



## Pengembangan Instrumen Literasi Sains Kontekstual Bencana Alam (Gempa Bumi dan Gunung Api)



Alimah Nuryanti<sup>1,\*</sup>, Ida Kaniawati<sup>2</sup>, Irma Rahma Suwarma<sup>2</sup>

<sup>1</sup>Program Studi PGMI UIN Raden Intan Lampung

<sup>2</sup>Program Studi Pendidikan Fisika Universitas Pendidikan Indonesia

\*Email: alimahnuryanti@radenintan.ac.id

DOI: <https://doi.org/10.33369/pendipa.9.3.718-725>

### ABSTRACT

*This study aimed to develop a valid and reliable scientific literacy test instrument for seventh-grade junior high school students on earthquakes and volcanoes. The method used is Research and Development with a 4D model, which includes the stages of define, design, develop, and disseminate. The research subjects consisted of 30 junior high school students. The development data were analyzed to assess the validity, reliability, difficulty level, and discriminatory power of the instrument. The development procedure included planning, preparation of the initial draft, validation of content and construct by experts through the suitability of scientific literacy indicators with the test items, followed by field trials, revisions, and implementation of the instrument. Based on 15 test items prepared, 13 were declared valid, one was revised, and one was not used. The analysis results showed the reliability of the instrument with a KR-20 coefficient of 0.76 which is included in the high category. The level of difficulty of the questions consisted of 12 items in the medium category and 3 items in the easy category, while the discriminatory power of all questions was in the very good to good category.*

**Keywords:** *Development of test instruments; scientific literacy; earthquakes; volcanoes.*

### ABSTRAK

Penelitian ini bertujuan mengembangkan instrumen tes literasi sains yang valid dan reliabel untuk siswa kelas VII SMP pada materi gempa bumi dan gunung api. Metode yang digunakan adalah Research and Development dengan model 4D, yang mencakup tahap define (pendefinisian), design (perancangan), develop (pengembangan), dan disseminate (penyebarluasan). Subjek penelitian terdiri atas 30 siswa SMP. Data hasil pengembangan dianalisis untuk menilai validitas, reliabilitas, tingkat kesukaran, serta daya pembeda instrumen. Prosedur pengembangan meliputi perencanaan, penyusunan draf awal, validasi isi dan konstruk oleh pakar melalui kesesuaian indikator literasi sains dengan butir soal, diikuti uji coba lapangan, revisi, dan implementasi instrumen. Dari 15 butir soal yang disusun, 13 soal dinyatakan valid, satu soal direvisi, dan satu soal tidak digunakan. Hasil analisis menunjukkan reliabilitas instrumen dengan koefisien KR-20 sebesar 0,76 yang termasuk kategori tinggi. Tingkat kesukaran soal terdiri dari 12 butir berkategori sedang dan 3 butir berkategori mudah, sedangkan daya pembeda seluruh soal berada pada kategori sangat baik hingga baik.

**Kata kunci:** pengembangan instrumen tes; literasi sains; gempa bumi; gunung api.

### PENDAHULUAN

Tingkat literasi sains siswa Indonesia hingga kini masih menjadi persoalan serius. Data PISA 2022 menunjukkan bahwa skor rata-rata literasi sains pelajar Indonesia masih jauh tertinggal dari rata-rata negara OECD, yang menandakan kesulitan siswa dalam menghubungkan konsep sains dengan persoalan

kehidupan sehari-hari (OECD, 2023). Hasil serupa juga teridentifikasi melalui beberapa kajian dalam negeri, yang menegaskan bahwa mayoritas siswa belum memiliki kemampuan literasi sains yang memadai untuk memahami isu-isu ilmiah yang relevan dengan kehidupan masyarakat modern (Rediani dkk., 2024; Zulaiha & Mulyono, 2025). Fakta ini memperlihatkan

perlu pbenahan serius pada kurikulum, strategi pembelajaran, sekaligus sistem evaluasi yang digunakan di sekolah.

Sejumlah faktor diketahui berkontribusi terhadap rendahnya capaian tersebut. Selain keterbatasan sarana belajar dan metode pengajaran yang masih berorientasi pada hafalan, persoalan mendasar lain adalah minimnya instrumen penilaian yang secara khusus dirancang untuk mengukur literasi sains secara utuh. Instrumen yang digunakan selama ini cenderung bersifat umum, tidak cukup menekankan dimensi proses dan konteks, serta jarang menyesuaikan dengan fenomena nyata yang dihadapi siswa (Zhang dkk., 2023; Andaresta, 2023). Akibatnya, profil kemampuan literasi sains yang diperoleh dari asesmen di sekolah sering kali kurang representatif, bahkan tidak mencerminkan kemampuan sebenarnya ketika siswa dihadapkan pada persoalan yang dekat dengan kehidupan sehari-hari, seperti fenomena kebencanaan (Susiani dkk., 2022; Alfi dkk., 2024).

Salah satu strategi yang dapat dilakukan adalah mengembangkan instrumen pengukuran berbasis kerangka PISA yang diadaptasi sesuai konteks lokal. Kerangka ini menekankan tiga aspek utama, yakni kemampuan menjelaskan fenomena secara ilmiah, mengevaluasi bukti, serta menggunakan sains dalam pemecahan masalah (Zhang dkk., 2023; Purnomo dkk., 2023). Dengan memasukkan topik yang relevan, seperti gempa bumi dan gunung api, instrumen tidak hanya berfungsi sebagai alat evaluasi kemampuan konseptual, tetapi juga mampu menilai keterampilan berpikir kritis dan penalaran siswa dalam menghadapi persoalan nyata. Beberapa penelitian sebelumnya juga membuktikan bahwa instrumen yang dikembangkan dengan acuan PISA dan divalidasi melalui uji coba lapangan dapat memberikan hasil pengukuran yang lebih sahih dan reliabel (Valio dkk., 2025).

Pemilihan topik gempa bumi dan gunung api sebagai konteks instrumen ini sangat beralasan. Indonesia yang berada di jalur *Ring of Fire* merupakan salah satu negara dengan tingkat kerawanan bencana geologi tertinggi di dunia. Banyak daerah di tanah air yang langsung bersinggungan dengan risiko letusan gunung api maupun gempa tektonik (Wardhani dkk., 2024;

Hidayat, 2024). Dengan mengaitkan soal-soal instrumen pada realitas tersebut, siswa tidak hanya diajak memahami teori geosains, tetapi juga belajar bagaimana mengaitkan konsep ilmiah dengan strategi mitigasi dan keputusan yang berhubungan langsung dengan kehidupan masyarakat. Konteks semacam ini terbukti meningkatkan motivasi belajar, relevansi, sekaligus kesadaran kesiapsiagaan bencana pada siswa (Wardhani dkk., 2024; Hidayat, 2024).

Manfaat dari pengembangan instrumen literasi sains ini tidak hanya dirasakan pada level praktis, melainkan juga akademis. Secara praktis, guru memperoleh alat ukur yang valid dan reliabel untuk menilai sejauh mana pemahaman sains siswa pada topik kebencanaan. Secara akademis, instrumen ini dapat dijadikan sarana pretest dan posttest dalam mengevaluasi efektivitas berbagai inovasi pembelajaran, seperti pendekatan STEM maupun *problem-based learning*. Sejumlah studi menunjukkan bahwa intervensi pembelajaran berbasis STEM berpotensi signifikan dalam meningkatkan kemampuan literasi sains dan keterampilan berpikir kritis, asalkan disertai dengan alat ukur yang sesuai (Nilyani, dkk., 2023; Purnomo dkk., 2023). Lebih jauh, hasil pengukuran yang akurat dapat menjadi rujukan penting bagi guru, peneliti, maupun pengambil kebijakan untuk merancang strategi pembelajaran sains yang lebih tepat bagi siswa Indonesia.

## METODE PENELITIAN

Penelitian ini menerapkan metode *Research and Development* (R&D) dengan model pengembangan 4D yang diperkenalkan oleh Thiagarajan, Semmel, dan Semmel (1974), yang mencakup tahapan *define, design, develop, dan disseminate*. Pada tahap *define*, peneliti melakukan pengumpulan informasi melalui kajian pustaka, analisis kurikulum, serta telaah hasil penelitian terdahulu yang relevan dengan pengembangan instrumen literasi sains. Tahap *design* difokuskan pada penyusunan kisi-kisi instrumen berdasarkan indikator literasi sains yang selaras dengan pendekatan pembelajaran STEM pada topik gempa bumi dan gunung api. Selanjutnya, tahap *develop* dilakukan dengan merumuskan butir instrumen sesuai kisi-kisi yang telah dirancang, kemudian melalui proses validasi isi dan konstruk oleh pakar di bidang

pendidikan sains. Pada tahap *disseminate*, instrumen yang telah dikembangkan diujicobakan secara terbatas pada 30 siswa SMP di wilayah Jawa Barat. Analisis kuantitatif terhadap data hasil uji coba mencakup uji validitas, reliabilitas, tingkat kesukaran, serta daya pembeda untuk menilai kelayakan instrumen. Berdasarkan hasil analisis, peneliti melakukan revisi instrumen dan mengimplementasikannya lebih lanjut guna mengukur literasi sains siswa setelah mengikuti pembelajaran STEM pada materi gempa bumi dan gunung api.

**HASIL DAN PEMBAHASAN**

**Tahap Define (Pendefinisian)**

Tahap pendefinisian menjadi fondasi awal dalam penyusunan instrumen literasi sains. Pada tahap ini, peneliti melakukan analisis kurikulum IPA SMP, khususnya materi gempa bumi dan gunung berapi. Hasil analisis menunjukkan bahwa topik kebencanaan ini tidak hanya memuat konsep-konsep ilmiah yang mendasar, tetapi juga memiliki keterkaitan erat dengan kehidupan sehari-hari siswa. Hal ini menjadikan materi gempa bumi dan gunung berapi relevan untuk digunakan sebagai konteks dalam mengukur literasi sains. Dengan demikian, kebutuhan akan instrumen asesmen literasi sains yang kontekstual dapat dirumuskan secara jelas, yakni instrumen yang mampu menilai

kemampuan siswa dalam memahami fenomena alam serta menghubungkannya dengan konsep ilmiah yang dipelajari di sekolah.

**Tahap Design (Perencanaan)**

Pada tahap perencanaan, konsep yang telah dirumuskan kemudian dikembangkan menjadi rancangan instrumen konkret. Peneliti menyusun kisi-kisi instrumen berbasis indikator literasi sains dengan mengacu pada *framework* PISA 2022 yang diterbitkan oleh OECD. Fokus pengukuran diarahkan pada aspek kompetensi, yaitu: (1) menjelaskan fenomena secara ilmiah, (2) mengevaluasi dan merancang penyelidikan ilmiah, serta (3) menginterpretasikan data dan bukti ilmiah. Untuk mengukur ketiga kompetensi tersebut, dirancang instrumen dalam bentuk tes pilihan ganda sebanyak 15 butir soal. Setiap butir soal dirancang berdasarkan indikator literasi sains yang relevan dengan konteks kebencanaan, sehingga selain mengukur pengetahuan konseptual, instrumen ini juga menuntut kemampuan siswa dalam berpikir kritis dan mengaitkan pengetahuan dengan kehidupan nyata. Rincian indikator literasi sains yang digunakan dapat dilihat pada Tabel 1.

**Tahap Develop (Pengembangan)**

Tahap pengembangan merupakan fase penting untuk memastikan kualitas instrumen yang telah dirancang. Setelah penyusunan butir

**Tabel 1.** Indikator Literasi Sains dalam Soal

Kompetensi Literasi Sains	Indikator Soal	No Soal
Menjelaskan fenomena ilmiah	Menjelaskan peristiwa pergerakan benua	3
	Membuat prediksi dampak pergerakan lempeng	4
	Menjelaskan faktor yang mempengaruhi besar letusan gunung api	11
	Menjelaskan penyebab peningkatan aktivitas vulkanik	12
	Memprediksi pengaruh faktor jarak pusat gempa terhadap ketahanan bangunan	13
Menginterpretasi data dan bukti ilmiah	Menjelaskan karakteristik lapisan bumi	1,2
	Menginterpretasi peta pergerakan tepi lempeng	5
	Menjelaskan terjadinya gempa bumi dengan menginterpretasi diagram pergerakan lempeng	6
	Menginterpretasi data sebaran gempa bumi dan gunung api	7
	Menginterpretasi hubungan antara aktivitas seismik dengan vulkanik	8
Mengidentifikasi dan mendesain penyelidikan ilmiah	Menjelaskan peristiwa pembentukan gunung api	9
	Menjelaskan dampak letusan gunung api	10
Mengidentifikasi dan mendesain penyelidikan ilmiah	Mengidentifikasi faktor yang mempengaruhi letusan gunung api	14,15

soal selesai, instrumen kemudian melalui proses validasi isi dan konstruk dengan melibatkan penilaian (*judgment*) dari para pakar. Validasi dilakukan dengan cara membandingkan kesesuaian antara indikator literasi sains yang ditetapkan dengan butir soal yang disusun. Para validator memberikan masukan terkait kesesuaian isi instrumen, apakah telah sesuai dengan indikator, masih memerlukan perbaikan, atau perlu dilakukan penggantian. Dalam penelitian ini, proses validasi melibatkan empat orang ahli di bidang pendidikan sains yang bertindak sebagai validator. Hasil telaah para ahli menunjukkan bahwa instrumen tes literasi sains yang dikembangkan telah memenuhi kriteria validitas isi dan konstruk.

Secara umum, hasil validasi menunjukkan bahwa kelima belas butir soal dinilai sesuai dengan konten dan indikator yang telah dirumuskan. Hal ini berarti setiap soal sudah mampu merepresentasikan aspek literasi sains yang ingin dicapai, baik pada dimensi menjelaskan fenomena ilmiah, merancang penyelidikan, maupun menginterpretasikan data dan bukti ilmiah.

Meskipun demikian, para validator juga memberikan sejumlah catatan perbaikan yang bersifat teknis maupun substantif. Beberapa soal memerlukan revisi minor, seperti perbaikan kualitas gambar agar lebih jelas (soal nomor 5 dan 9), penyesuaian redaksi opsi jawaban (soal nomor 6 dan 8), serta penyempurnaan uraian atau data yang digunakan dalam soal (soal nomor 11 dan 13). Pada soal nomor 5, validator menekankan perlunya memperjelas garis pada gambar serta melakukan revisi pada kunci jawaban. Sementara itu, soal nomor 6 disarankan untuk memperbaiki alternatif jawaban agar lebih proporsional. Catatan lain muncul pada soal nomor 8, di mana opsi D diperbaiki menjadi redaksi “terdapat hubungan...,” sehingga maknanya lebih akurat. Adapun soal nomor 9 dan 11 mendapatkan masukan berupa perbaikan ukuran gambar dan penyesuaian data dalam uraian soal.

Berikut adalah contoh soal literasi sains setelah dilakukan perbaikan sesuai catatan validator.

#### Contoh Soal 1

Indikator literasi sains:

Menjelaskan fenomena ilmiah

Alfred Wegener mengemukakan teori pergeseran benua berdasarkan adanya kesamaan fosil dan lapisan batuan antara pantai timur Amerika Selatan dan pantai barat Afrika. Temuan ini menunjukkan bahwa kedua benua tersebut dahulu pernah menyatu. Namun, teori tersebut sempat ditolak karena Wegener tidak dapat menjelaskan mekanisme pergerakan benua.

Menurut pendapatmu, apa faktor utama yang menyebabkan lempeng benua dapat bergerak?

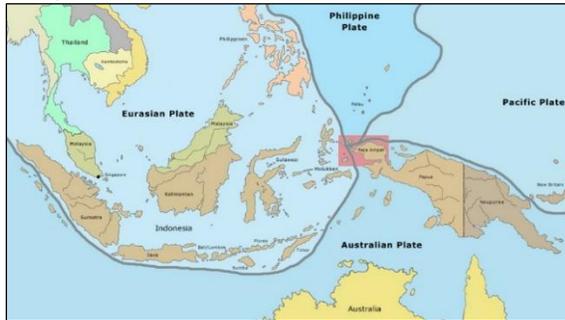
- a. Kerak bumi tersusun atas lempengan yang dapat bergerak bebas tanpa hambatan.
- b. Putaran inti bumi secara langsung mendorong lapisan kerak bumi untuk bergerak.
- c. Proses konveksi di dalam mantel bumi menyebabkan pergeseran lempeng yang mengapung di atas astenosfer yang plastis.
- d. Perubahan arah medan magnet bumi menggerakkan daratan benua.

Soal tentang pergerakan benua ini dirancang untuk mengukur literasi sains pada kompetensi menjelaskan fenomena ilmiah melalui tiga aspek utama. Pertama, siswa diminta memahami mekanisme ilmiah di balik teori pergeseran benua Wegener, yaitu peran arus konveksi di mantel bumi dalam menggerakkan lempeng di atas astenosfer plastis. Kedua, kemampuan menjelaskan fenomena nyata diuji dengan menghubungkan konsep ilmiah tentang dinamika bumi dengan peristiwa geologi yang ada. Ketiga, keberadaan opsi jawaban yang merepresentasikan miskonsepsi umum berfungsi untuk menilai sejauh mana siswa mampu membedakan penjelasan ilmiah yang benar dari penalaran yang tidak sesuai sains.

## Contoh Soal 2

Indikator literasi sains:

Menjelaskan fenomena ilmiah



**Gambar 1.** Lempeng Benua

Berdasarkan Gambar 1, Indonesia berada pada titik pertemuan tiga lempeng tektonik yaitu lempeng Australia, lempeng pasifik, dan lempeng Eurasia. Hal ini menyebabkan munculnya deretan gunung api di Indonesia. Australia juga merupakan negara yang berdekatan dengan pertemuan lempeng Australia dan Erurasia, namun tidak muncul deretan gunung aktif seperti halnya Indonesia. Hal ini dikarenakan...

- Lempeng tektonik yang berbatasan dengan Australia tidak sebanyak di Indonesia
- Gunung hanya terbentuk pada 3 lempeng yang saling bertemu
- Australia sebenarnya berada pada lempeng Australia saja
- Wilayah Australia bukan merupakan daerah tepi lempeng tempat menunjammnya lempeng-lempeng yang saling bertemu

Soal ini menilai kompetensi literasi sains dalam aspek menginterpretasikan data dan bukti ilmiah melalui konteks pertemuan lempeng tektonik. Siswa diminta menganalisis informasi dari peta/gambar mengenai posisi Indonesia dan Australia terhadap lempeng tektonik, lalu menghubungkannya dengan fenomena geologi

berupa keberadaan gunung api. Jawaban yang benar menunjukkan kemampuan siswa menginterpretasikan bukti ilmiah bahwa aktivitas gunung api terbentuk pada zona subduksi, seperti di Indonesia, sedangkan Australia tidak berada di tepi lempeng yang menunjam, sehingga tidak muncul deretan gunung api aktif.

### Tahap Disseminate

Tahap disseminate merupakan proses uji coba soal kepada subjek penelitian. Dari uji coba tersebut dilakukan analisis validitas butir soal secara statistik. Pada penelitian ini, validitas butir soal dianalisis menggunakan korelasi point biserial antara skor tiap butir soal dengan skor total. Nilai korelasi ( $r$ ) yang diperoleh kemudian diinterpretasikan, dan soal dinyatakan valid apabila nilai  $r$  lebih tinggi dari standar minimal uji validitas. Selanjutnya, reliabilitas instrumen dihitung dengan menggunakan rumus KR-20 karena tes berbentuk objektif dengan data dikotomi. Hasil perhitungan menunjukkan nilai reliabilitas sebesar 0,76 yang masuk kategori tinggi. Selain itu, dilakukan pula analisis tingkat kesukaran dan daya pembeda soal. Rekapitulasi hasil uji validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal disajikan pada Tabel 2.

Hasil analisis validitas menunjukkan bahwa dari 15 butir soal yang dikembangkan, 13 butir dinyatakan valid dengan nilai korelasi point biserial di atas 0,30, sementara 2 butir (nomor 12 dan 15) tidak memenuhi kriteria karena memiliki korelasi rendah, masing-masing sebesar 0,117 dan 0,178. Temuan ini menegaskan bahwa sebagian besar butir soal telah konsisten dengan konstruk literasi sains yang diukur, sesuai dengan pendapat Azwar (2021) bahwa instrumen dapat dinyatakan valid apabila butir soal memiliki korelasi positif yang signifikan dengan skor total.

**Tabel 2.** Rekapitulasi hasil uji validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal

No	Validitas Butir		Tingkat Kesukaran		Daya Beda		Keterangan
	r	Kategori	Nilai	Kategori	Nilai	Kategori	
1	0,482	valid	0,621	Sedang	0,625	Baik	Dipakai
2	0,502	valid	0,655	Sedang	0,500	Baik	Dipakai
3	0,548	valid	0,586	Sedang	0,750	Baik Sekali	Dipakai
4	0,624	valid	0,517	Sedang	0,875	Baik Sekali	Dipakai
5	0,523	valid	0,655	Sedang	0,625	Baik	Dipakai
6	0,461	valid	0,414	Sedang	0,625	Baik	Dipakai
7	0,505	valid	0,517	Sedang	0,500	Baik	Dipakai
8	0,559	valid	0,793	Mudah	0,500	Baik	Dipakai
9	0,481	valid	0,655	Sedang	0,625	Baik	Dipakai
10	0,482	valid	0,620	Sedang	0,750	Baik Sekali	Dipakai
11	0,649	valid	0,621	Sedang	0,750	Baik Sekali	Dipakai
12	0,117	Tidak valid	0,655	Sedang	0	Jelek	Tidak Dipakai
13	0,518	valid	0,758	Mudah	0,500	Baik	Dipakai
14	0,422	valid	0,896	Mudah	0,250	Cukup	Dipakai
15	0,178	Tidak valid	0,586	Sedang	0,375	Cukup	Revisi

Reliabilitas : 0,76

Hasil ini menunjukkan bahwa indikator literasi sains berbasis PISA 2022 dapat diterjemahkan secara baik dalam konteks kebencanaan, khususnya gempa bumi dan gunung api. Sejalan dengan Bybee (2020), instrumen yang dikembangkan dengan konteks autentik akan lebih mudah dipahami siswa. Reliabilitas instrumen dihitung dengan KR-20 menghasilkan nilai sebesar 0,76, termasuk kategori tinggi, yang menunjukkan konsistensi instrumen dalam mengukur literasi sains siswa (Arikunto, 2021). Temuan ini didukung oleh penelitian Liu dkk. (2023) dan Zhang dkk., (2024) yang menunjukkan bahwa instrumen berbasis konteks autentik cenderung memiliki reliabilitas yang lebih stabil

Analisis tingkat kesukaran memperlihatkan bahwa mayoritas butir berada pada kategori sedang (12 soal), sementara 3 soal berada pada kategori mudah. Kondisi ini menunjukkan bahwa instrumen tidak terlalu sulit maupun terlalu mudah, sehingga sesuai digunakan untuk mengukur kemampuan literasi sains siswa secara merata. Menurut Sudjana (2020), tingkat kesukaran moderat ideal karena dapat membedakan kemampuan siswa tinggi, sedang, dan rendah. Pada aspek daya pembeda, sebagian

besar soal berada dalam kategori baik hingga sangat baik, dengan nilai 0,50–0,875. Hal ini menandakan bahwa soal mampu mengidentifikasi perbedaan kemampuan siswa secara efektif. Hasil ini sejalan dengan Arikunto (2021) serta penelitian kontemporer oleh Susanti dkk., (2023) dan Rachmawati dan Wulandari (2022) yang menegaskan bahwa instrumen berkualitas harus memiliki daya pembeda yang kuat agar dapat berfungsi optimal sebagai alat ukur.

Konteks kebencanaan yang digunakan dalam instrumen ini, yaitu gempa bumi dan gunung api, memberikan nuansa autentik dalam asesmen. Pendekatan berbasis konteks terbukti dapat meningkatkan literasi sains siswa (Kartini dkk., 2022), memperkuat keterampilan pemecahan masalah (Hikmawati dkk., 2022), sekaligus membangun kesiapsiagaan menghadapi bencana (Novia dkk., 2021; Rahman dkk., 2022). Instrumen ini juga disusun berdasarkan prinsip Evidence-Centered Design (ECD), di mana setiap butir soal merepresentasikan bukti nyata dari kompetensi yang diukur (Pellegrino dkk., 2021). Dengan demikian, validitas konseptual instrumen semakin kuat. Secara lebih luas, pengembangan

instrumen ini sejalan dengan tren asesmen global seperti PISA 2022 yang menekankan pentingnya keterkaitan antara sains dengan isu nyata (OECD, 2022). Lebih lanjut, asesmen berbasis kinerja dan konteks autentik terbukti dapat memperkuat kreativitas dan inovasi siswa dalam pembelajaran STEM (Akdemir-Beveridge dkk., 2025).

## KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa pengembangan instrumen literasi sains berbasis konteks kebencanaan gempa bumi dan gunung api melalui model 4D berhasil menghasilkan instrumen yang valid, reliabel, serta layak digunakan dalam pembelajaran sains. Dari 15 butir soal yang disusun, 13 dinyatakan valid dengan nilai korelasi di atas 0,30, sementara reliabilitas yang diperoleh sebesar 0,76 menunjukkan konsistensi tinggi instrumen. Analisis tingkat kesukaran memperlihatkan sebagian besar soal berada pada kategori sedang dengan daya pembeda baik hingga sangat baik, yang menandakan kemampuan instrumen dalam membedakan variasi kemampuan siswa. Temuan ini membuktikan bahwa penerapan konteks autentik kebencanaan tidak hanya meningkatkan relevansi soal, tetapi juga memperkuat validitas konseptual instrumen. Secara praktis, guru memperoleh alat ukur yang dapat diandalkan untuk menilai literasi sains siswa, sementara secara akademis instrumen ini dapat dimanfaatkan dalam penelitian maupun evaluasi pembelajaran berbasis STEM. Dengan demikian, instrumen yang dikembangkan tidak hanya berkontribusi pada peningkatan kualitas asesmen literasi sains di Indonesia, tetapi juga mendukung upaya global untuk membangun keterampilan abad ke-21 melalui integrasi sains dengan isu nyata kehidupan.

## DAFTAR PUSTAKA

Akdemir-Beveridge, dkk. (2025). Performance-based assessments for STEM literacy and creativity. *Journal of Science Education and Technology*, 34(2), 215–232.

Andaresta, O. (2023). Development of assessments to measure students' science literacy. *Semantic Scholar*. <https://www.semanticscholar.org>

Arikunto, S. (2021). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.

Azwar, S. (2021). *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar.

Bybee, R. (2020). *Science education and context-based assessment*. London: Routledge.

Hidayat, Z. (2024). The role of education and training for disaster-affected communities. *Atlantis Press*. <https://www.atlantispress.com>

Hikmawati, N., Astuti, S., & Rahma, Y. (2022). STEM-based disaster context learning to improve problem-solving. *International Journal of Instruction*, 15(3), 227–244.

Kartini, E., Suryana, A., & Fajri, M. (2022). Contextual disaster-based learning to enhance science literacy. *Jurnal Pendidikan IPA Indonesia*, 11(1), 25–36.

Liu, Y., Chen, X., & Wang, H. (2023). Developing reliable science literacy assessments with authentic contexts. *Journal of Science Education Research*, 32(4), 451–466.

Nilyani, K., Asrizal, A., & Usmeldi, U. (2023). The effect of STEM integrated science learning on scientific literacy and critical thinking skills of students: A meta-analysis. *Jurnal Penelitian Pendidikan IPA*, 9(6), 65–72. <https://doi.org/10.29303/jppipa.v9i6.261>

Novia, L., Hidayat, A., & Rafi, M. (2021). Assessment of disaster literacy among junior high school students. *Journal of Disaster Education Research*, 7(2), 55–64.

OECD. (2022). *PISA 2022 assessment and analytical framework*. Paris: OECD Publishing.

OECD. (2023). *PISA 2022 results: Country note—Indonesia*. OECD. <https://www.oecd.org>

Pellegrino, J., Chudowsky, L., & Glaser, R. (2021). *Evidence-centered design in science assessment* (Updated ed.). Washington, DC: National Academies Press.

Purnomo, S., dkk. (2023). Effectiveness of ADI-STEM to improve student's science literacy. *International Journal of Recent Educational Research (IJORER)*. <https://journal.ia-education.com>

- Rachmawati, D., & Wulandari, F. (2022). Pengembangan instrumen literasi sains berbasis konteks. *Jurnal Inovasi Pendidikan IPA*, 4(1), 15–27.
- Rediani, D., dkk. (2024). Scientific literacy in Indonesian secondary education. *Indonesian Journal of Educational Development*. <https://ojs.mahadewa.ac.id>
- Susiani, K., Dantes, N., Arnyana, I. B. P., & Suarni, N. K. (2022). Developing testing instruments to measure science literacy of elementary school students. *Cypriot Journal of Educational Sciences*, 17(12), 4361–4378. <https://doi.org/10.18844/cjes.v17i12.773>
- Alfi, B. R., Hw, P., & Pratama, A. (2024). Validating instrument to assess science literacy and independent learning skills in high school. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, 10(3), 979–988. <https://doi.org/10.22219/jpbi.v10i3.34770>
- Sudjana, N. (2020). *Penilaian hasil proses belajar mengajar*. Bandung: Remaja Rosdakarya.
- Susanti, R., Nugroho, P., & Kurniawati, S. (2023). Item analysis of STEM literacy assessment instruments. *Jurnal Pendidikan Sains Indonesia*, 9(2), 143–153.
- Valio, F. A., dkk. (2025). Development of a science literacy test for junior high (PISA-based). *Jurnal Pendidikan Matematika dan IPA (JPMIPA)*, Universitas Lampung. <https://jpmipa.fkip.unila.ac.id>
- Wardhani, P. I., dkk. (2024). Evaluation of disaster safe education unit programme (SPAB) — Merapi. *PubMed Central (PMC)*. <https://www.ncbi.nlm.nih.gov/pmc>
- Zhang, L., dkk. (2023). Development and validation of an instrument for assessing science literacy. *International Journal of STEM Education*. <https://stemeducationjournal.springeropen.com>
- Zhang, L., Huang, J., & Li, M. (2024). Reliability of contextualized science assessment instruments. *International Journal of STEM Education*, 10(1), 56–72.
- Zulaiha, J., & Mulyono, M. (2025). Condition of science literacy in Indonesia. *Jurnal Pendidikan IPA Indonesia (JPII)*. <https://journal.unnes.ac.id>